DOCUMENT RESUME

ED 221 589 TM 820 626

TITLE Interpreting and Using Commercial Achievement Test Results. Basic Skills Improvement Policy: Supplement to Implementation Guide 1.

INSTITUTION Massachusetts State Dept. of Education, Boston. Bureau of Research and Assessment.

PUB DATE May 82
NOTE 40p.

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Achievement Tests; *Basic Skills; Elementary Secondary Education; Listening Comprehension Tests; Quantitative Tests; Reading Tests; Scores; Speech Tests; *Standardized Tests; Student Evaluation; Test Interpretation; *Test Results; Test Use; Writing Evaluation
IDENTIFIERS *Basic Skills Improvement Policy (Massachusetts)

ABSTRACT

This supplement is intended to help schools get the maximum amount of useful information from test results; information that will be helpful in program monitoring, grouping, planning instructional activities, and reporting results to parents and community. Contents include an introduction to test interpretation and interpreting of test scores; suggestions for using information in addition to test scores to improve the quality of decisions concerning students; suggestions for using test data for monitoring performance at the district level and to identify instructional needs; and guidelines for working with the press. A chart for interpreting normal curve equivalents, percentiles, stanines and standard deviation units, and sample press releases are appended. (PN)

**********
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
**********

ED221589

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

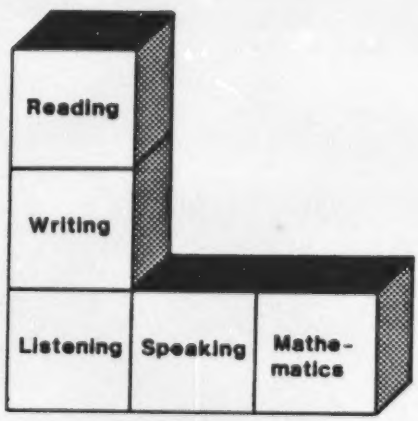
X This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

## Basic Skills Improvement Policy

# Supplement to Implementation Guide #1

# Interpreting and Using Commercial Achievement Test Results




"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
C. D. Bills
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

TM 820 626

MASSACHUSETTS DEPARTMENT OF EDUCATION

MAY 1982

MASSACHUSETTS BOARD OF EDUCATION

Donald R. Walker, Chairperson

Howard A. Greis, Vice Chairperson

John W. Bond
Millie Clements
Joseph P. Crimmins
James R. Grande
James L. Green
Charles T. Grigsby
Armando Martinez
Edwin Rossman
Joseph C. Savery
Dorothea Zanetti

John H. Lawson, Commissioner of Education
John B. Duff, Ex. Officio, Chancellor, Board of Regents

BUREAU OF RESEARCH AND ASSESSMENT

Allan S. Hartman, Director
Matthew H. Towle
Dorothy F. Earle
Richards S. Williams
Betty Hancock

Publication # 12822 approved by John Manton, State Purchasing Agent.

The Massachusetts Department of Education insures equal employment/ educational opportunities/and affirmative action, regardless of race, color, creed, national origin, or sex, in compliance with Title IX, or handicap, in compliance with Section 504.

# BASIC SKILLS IMPROVEMENT POLICY

Supplement

to

Implementation Guide #1

Interpreting and Using Commercial Achievement

Test Results

Bureau of Research and Assessment
Department of Education
Quincy Center Plaza
1385 Hancock Street
Quincy, Massachusetts 02169

May 1982

Copyright © 1982 Massachusetts Department of Education

This publication is restricted for use in Massachusetts schools for the specific purpose of implementing the Massachusetts Basic Skills Improvement Policy and Regulations and may be duplicated by Massachusetts schools for that purpose. Duplication for use for other purposes requires written permission from the Massachusetts Department of Education.

## FOREWORD

Between 70 and 80 percent of Massachusetts school systems are using commercially available tests to evaluate students' basic skill achievement in elementary level reading and mathematics. Approximately 19 percent use such tests to evaluate student achievements at the secondary level. Between thirty and forty different commercial tests are used including: the California Achievement Test; the Iowa Test of Basic Skills'; the Comprehensive Test of Basic Skills; the Stanford Achievement Test; the Metropolitan Achievement Test; the SRA Achievement Series, and the Stanford Diagnostic Test.

While it is clear that the major use of these tests, in the context of the Policy, is to determine performance in terms of local standards, the test results can be used for other purposes as well.

The purpose of this supplement is to help schools get the maximum amount of useful information from test results -- information that will be helpful in: program monitoring; grouping; planning instructional activities; and reporting results to parents and community. As such this information should be useful to:

- The teacher who is attempting to determine instructional needs;
- The principal who is attempting to monitor instructional activity;
- The curriculum specialist who is attempting to focus and plan curriculum improvement efforts;
- The district administrator who is attempting to report accurately and clearly to the community.

## ACKNOWLEDGEMENTS

Materials on using test data at the district level have been drawn in part from the Handbook for Proficiency Assessment, (Sacramento, California: California State Department of Education, 1979). Chapter IV of this Supplement, "Using Test Data to Identify Instructional Needs," is derived from Using and Interpreting Results of Standardized Achievement Tests. This publication is part of Educational Testing Service's series on the uses of measurement.

CONTENTS


| | | Page |
|---|---|---|
| I. | INTRODUCTION TO TEST INTERPRETATION | 1 |
| | Standardized Achievement Tests | 1 |
| | Norm - Referenced Tests | 1 |
| | Test Scores that You will be Using | 2 |
| | Criterion and Objectives-Referenced Tests | 3 |
| | Some Pointers for Interpreting Test Scores | 4 |
| II. | DECISIONS ABOUT STUDENTS: USING INFORMATION IN ADDITION TO TEST SCORES | 5 |
| | Multiple Measures of Performance | 5 |
| | Why Test Scores Should Not be Used as Exclusive Measures | 5 |
| | Locating Other Sources of Information | 5 |
| | Types of Additional Information | 6 |
| | Collecting New Information | 6 |
| III. | USING TEST DATA FOR MONITORING PERFORMANCE AT THE DISTRICT LEVEL | 8 |
| | Considerations at the District Level | 8 |
| | Combining Test Scores | 8 |
| | Determining Group Averages | 9 |
| | Group Averages: What to Look For | 9 |
| | Rules of Thumb for Interpreting Differences | 9 |
| IV. | USING TEST DATA TO IDENTIFY INSTRUCTIONAL NEEDS | 15 |
| | Interpreting the Individual Item Analysis Report | 15 |
| | Looking at Group Results | 15 |
| | Looking at Domain Results | 17 |
| | Looking at Individual Results | 17 |
| | Planning Steps | 18 |
| V. | WORKING WITH THE PRESS | 19 |
| | Guidelines for Presenting Information to the Press | 19 |
| | Steps to Follow in Working with the Press | 20 |
| | Stage One -- Background | 20 |
| | Stage Two -- Getting Out the News | 23 |
| | Stage Three -- Follow-up | 24 |
| APPENDIX A | AREA UNDER THE NORMAL CURVE DIVIDED INTO NCEs, PERCENTILES, STANINES, AND STANDARD DEVIATION UNITS | A-1 |
| APPENDIX B | SAMPLE PRESS RELEASES | B-1 |

# I. INTRODUCTION TO TEST INTERPRETATION

## Standardized Achievement Tests

Most students take standardized achievement tests at some point during their school years, and the results are given in the form of test scores.

Achievement tests are made up of questions that represent a sample of the many questions that could be asked about a particular subject, or that would demonstrate a particular skill. The student's performance on this sample allows us to judge how well the student has learned the subject or skill.

There are two basic types of achievement tests; norm-referenced and criterion or objectives-referenced tests.

## Norm-Referenced Tests

Norm-referenced tests measure a student's knowledge and skill in relation to the performance of other students who have taken the same test under the same conditions. This lets us know where a student stands relative to other students in the same grade.

Questions that appear in standardized achievement tests are written by experts who are familiar with curricula currently taught throughout the country. Each question is tested with groups of students, and the results of these "pilot" tests are used to identify poor questions which are revised or discarded. The final versions of the test questions are then assembled into a total test, and the test is given to a large group of students, who are representative of students across the country. The test results of these large groups are used to establish the "norms" or standards to which an individual student's test performance is compared. These norms are not

judgments about what performance should be, but are instead averages of actual performance by a representatively selected group.

Sometimes norms are computed for other groups, such as students in the state or the local district or different socioeconomic groups. This allows comparison of students with other students like themselves.

Test Scores That You Will Be Using[1]

Here are some types of test scores that you receive from test publishers and which we refer to in the following sections of this supplement.

- Raw Scores - A raw score is the number of questions answered correctly.
- Percentile Score - Percentiles indicate the student's position relative to the norm group. Percentiles should not be confused with percentage scores. The latter give the percentage of questions answered correctly, while a percentile score represents the percent of the norm group that the individual student equalled or surpassed in his/her test performance. Percentile ranks are sometimes shown as ranges or bands in order to account for the amount of error inherent in any measurement.

---

EXAMPLE

A Fourth Grader has a National Percentile Rank of 52

This means that 52 percent of the pupils in the norm group scored equal to or lower than our sample fourth grader on the test. Since the norm group was representative of all fourth graders in the nation, it is estimated that the student performed as well or better than 52 percent of all the fourth graders in the nation. For most standardized tests, percentile ranks are developed separately for each grade and for a particular time of the year.

---

- Stanines - A stanine is a score on a nine-unit scale from 1 to 9, where a score of 5 describes average performance. The highest stanine is 9; the lowest is 1. A stanine score places a student on this scale by comparing his or her performance with the performance of a group of other students. As with the percentile score, this norm group can be made up of a national sample or a smaller local sample. Unless the norm group is described, the score has little meaning.

---

[1]Appendix A shows distributions of test scores and their correspondence with each other.

- Standard Score - A standard score is a general term applied to a transformed raw score. This adjustment is made so that scores on different test forms or scores for different levels can be compared.

- Grade Equivalents - A grade equivalent indicates the grade level, in years and months, for which a given score was the average or middle score in the norm group. For example, a score of 25 with the grade equivalent of 4.6 means that, in the norm group, 25 was the average score of students in the sixth month of the fourth grade. If another student in the sixth month of the fourth grade were to take the same fourth grade test and score 25 correct, his/her performance woud be "at grade level" or average for his/her grade placement. If he/she were to get 30 right, which might be a grade equivalent of 5.3, he/she would have done as well as the typical fifth grader in the third month of the school year on that fourth grade test. This does not mean that the fourth grader can do all fifth grade work. There are many things a fifth grader has learned that are not measured on a fourth grade test.

  Similarly, a 3.3 grade equivalent for a fourth grader would mean that he/she is performing, on the fourth grade test, the way the average pupil in the third month of third grade would perform on the same test. It does not suggest that he/she has learned only third grade material. When a student's grade equivalent is markedly higher or lower than the student's grade at the time of testing, the score should be interpreted with caution. Keep in mind that, since it is an average, half of all the students must score above, and half below, grade level.

  Although grade equivalents may sound like a simple idea, they can be easily misunderstood. For this reason, schools are increasingly coming to rely on percentiles and stanines as more valid ways of interpreting scores.

Criterion or Objectives-Referenced Tests

The State secondary level tests designed to measure students' basic skills in mathematics and reading are criterion or objectives-referenced tests. These tests measure a student's knowledge and skill against a "criterion" or specific objectives indicating what a student should know or be able to do as a result of the school's program of instruction. This kind of test report helps the teacher to plan instruction for the student or for groups of students, based on what they still need to learn.

Instead of comparing student performance to the performance of other groups of students, these tests report performance on specific skills. An example of such a skill is "subtraction of two 2-digit numbers involving borrowing." "Mastery" of this skill would be shown by meeting a standard established by the school district, such as answering correctly 85 percent of the questions that test this skill. The determination of what "mastery" means requires judgment on the part of those responsible for setting the standard.

SOME POINTERS FOR INTERPRETING TEST SCORES

- Be sure that the norms you're using for comparison purposes are appropriate. The norms group should be as much like your group as possible.
- Do not use raw scores to compare performance in different subjects or different test forms. Use stanines or another type of standard score for this purpose.
- If percentile bands on two tests in the larger test battery overlap, it is unlikely that the difference in the scores is significant or meaningful; conversely, if they don't overlap, the difference is probably significant and meaningful.
- Stanines should generally be more than one point apart to be considered meaningful. Generally, performance in stanines 1-3 is below average, 4-6 is average, and 7-9 is above average.

## II. DECISIONS ABOUT STUDENTS: USING INFORMATION IN ADDITION TO TEST SCORES

### Multiple Measures of Performance

The integration of test results with other information about student performance e.g., teacher ratings, can improve the quality of decisions we make about such things as student placement and grouping. The use of data in addition to test scores should always be encouraged because test scores have limitations even though they may be valid and reliable.

### Why Test Scores Should not be Used as Exclusive Measures

1. A test is no more than a school-type job. A test is a work sample, a sample of how well a student works with a given set of learning materials. No two tests are alike in the jobs they ask students to do, and no two test scores mean exactly the same thing.

2. A test score is an estimate. Test scores are sensible approximations. They are not, however, pin-point measurements; they should not be treated as exact numerical measures.

3. No test measures all of any ability. Tests are samplings of skills and abilities. Some tests cover only a narrow range of knowledge or skill, and no test can cover all the knowledge a person might have.

4. Performance on tests can change. It is not fair to judge a student's knowledge or skill on the basis of a single test score. Some students are poor test takers in general, and students are always changing -- learning new information and skills.

### Locating Other Sources of Information

Before you select additional sources of information to supplement your basic skills test results, review existing records in your district. Look at students' permanent record cards and teachers' class records to determine the types of information that are available.

As you look through existing records, think about the feasibility of collecting and summarizing the information that appears to be available. How much time will it take? What will it cost?

Finally, try to determine the validity and reliability of existing information. Ask yourself: How appropriate would each type of data be for making decisions about student performance. It is particularly important to identify data sources that will be available for almost all the students for whom you are making decisions. (Large amounts of missing data create a number of problems.)

## Types of Additional Information

In addition to the basic skills test you are using, you should be able to identify data sources similar to those listed below. Other examples will occur as you examine school records:

- Diagnostic test results
- Results of locally developed tests
- Information on reading level
- Information on math level
- Grade point average
- Number of prerequisite skills mastered
- Teacher ratings of student performance.

## Collecting New Information

If you are not satisfied with the data already available, you should collect new information. Before doing so, however, you should systematically weigh the advantages and disadvantages of collecting new information.

## Collecting New Information

### Advantages

- You can avoid past mistakes, e.g., plug holes for missing data.
- You have the opportunity to train staff in data collection, thereby improving reliability.
- You can obtain the type of information you really want rather than being satisfied with information just because it is available.

### Disadvantages

- There may be substantial costs, particularly of staff time.
- It may not be possible to ensure comparability of data sources if one set of data is collected at one point in time and another set is collected at a different time.
- Data you already have available may be more "authentic" than data that you deliberately collect for a particular purpose.

# III. USING TEST DATA FOR MONITORING PERFORMANCE AT THE DISTRICT LEVEL

## Considerations at the District Level

For purposes of curriculum monitoring and instruction planning, it is useful to think of test results in terms of different levels of analysis -- the district as a whole, the classroom, and the individual student. This chapter includes procedures for summarizing and interpreting test results at the district level. The chapter that follows covers classroom and individual applications.

District administrators need information to help in the allocation of staff and material resources as well as information to aid in curriculum review and modification. Basic skills test results can be summarized in ways that contribute to effective district-wide decision making.

## Combining Test Scores

Some test publishers provide district-wide summaries. However, if these scoring services are not available, or if you require additional breakdowns, you will have to develop your own summaries. You should always keep the following points in mind as you combine results for district-wide use:

1. Do not average (or total) test scores over different content areas.

2. Do not calculate averages using grade equivalents or percentile scores.

3. Use the following scores in computing averages: scaled scores, raw scores, or percent correct. (These scores can be translated back into grade equivalents or percentile ranks if desired.)

## Determining Group Averages

For each content area, obtain the average score (based on percent correct, raw scores or scaled scores) for each grade level and for each school in the district. Exhibit 1 (page 11) shows the results of a mathematics achievement test administered to students in grades 7 and 9. If you have scaled scores, such as those reported for standardized achievement tests, you can calculate averages and arrange them to show school and grade comparisons. Exhibit 2 (page 13) presents summaries of scaled scores.

## Group Averages: What to Look For

In interpreting test results such as those presented in Exhibits 1 and 2, there are four key things to look for:

1. Do the average scores increase from one grade to the next?
2. Are the differences between schools minimal?
3. Are there unexpected differences between grades at a particular school?
4. Are there unexpected results as a particular group of students moves from grade to grade?

## Rules of Thumb for Interpreting Differences

As you compare results, you will want to know whether the differences you observe are minor and unimportant, or important ones that you should explore further. The following guidelines will help you make this decision:

1. Compare each group (school) average to the district average for each grade. If the difference in averages is less than the value of the corresponding standard deviation, the difference should be considered minimal.

2. Differences that are more than twice as great as the value of the standard deviation for that grade level should be considered substantial.

3. The number of students will affect results. As a rule of thumb, if there are fewer than 25 scores for a particular group, conclusions about that group should be drawn with considerable caution.

4. In looking at students' progress over time, be sure you have comparable groups, either by excluding students for whom you do not have complete information or by making the necessary statistical adjustments.

EXHIBIT 1

Average Number of Correct Responses on
a Mathematics Test of 75 Questions

| | Grade 7 | | Grade 9 | |
|---|---|---|---|---|
| | Average Score | Number of Students | Average Score | Number of Students |
| Blue School | 50.2 | (95) | 65.4 | (49) |
| Green School | 56.2 | (44) | 63.1 | (48) |
| Yellow School | XX | (None) | 67.8 | (56) |
| District Average | 52.1 | (139) | 65.6 | (153) |
| Standard Deviation[a] | 4.6 | | 3.2 | |

[a]Standard deviation is a measure of the dispersion of scores about the mean. The smaller the standard deviation, the more scores tend to cluster about the mean.

EXAMPLE

Interpreting Results in Exhibit 1

Following the guidelines presented on pages 9 and 10, a number of conclusions might be drawn about the data presented in Exhibit 1.

1. First, as expected, the average score for 9th graders (65.6) is higher than the average score for 7th graders (52.1).

2. For the most part, the differences between schools appear to be inconsequential. Even the performance of 7th graders in the Green School (56.2), which is somewhat higher than the average 7th grade score (52.1), does not exceed the standard deviation. (The score difference is 4.1 and the standard deviation is 4.6).

3. For some reason, there is a greater difference between grade level scores at the Blue School (15.2 difference between grades 7 and 9) than at the Green School (6.9 difference between grades 7 and 9). This may simply mean that the groups are not comparable, but we ought to check out whether the score difference is an expected one. If it is not, we should investigate and then check item analysis data. (See Chapter II.)

EXHIBIT 2

Distribution of Standard Scores for Current Sixth Grade Compared with their Fourth and Fifth Grade Performance on Standardized Reading Test

| | Grade 4 (1980) | | Grade 5 (1981) | | Grade 6 (1982) | |
|---|---|---|---|---|---|---|
| | Average Standard Score | Number of Students | Average Standard Score | Number of Students | Average Standard Score | Number of Students |
| Blue School | 425 | 114 | 436 | 118 | 441 | 115 |
| Green School | 415 | 117 | 419 | 120 | 421 | 126 |
| Yellow School | 430 | 95 | 450 | 114 | 465 | 112 |
| District Average | 425 | | 433 | | 441 | |
| Standard Deviation | 8 | | 7 | | 6 | |

**EXAMPLE**

**Interpreting Results in Exhibit 2**

Following the guidelines presented on pages 9 and 10, a number of conclusions might be drawn about the data presented in Exhibit 2.

1. First, a number of comparisons can be made because the standard scores (ranging from 400-490) allows comparisons to be made between grade level, regardless of the test form taken. Secondly, a gain of 4 points from the 4th to 5th grade represents the same performance increase as a gain of 4 points from the 5th to 6th grade.

2. If we focus on the current 6th grade, we see that the Green School is substantially below the district average (more than two standard deviations, or 12 points) and that the Yellow School is substantially above the district average.

3. Assuming we have checked to be sure that the groups are comparable, we can easily observe that there is a significant problem with reading performance at the Green School. As youngsters progress from grade to grade, they are falling further and further behind the district as a whole. Certainly we should investigate further. Perhaps you should then do diagnostic testing, classroom observation, etc.

4. The current 6th grade at the Yellow School is progressing by leaps and bounds. Opportunities for advancement and enrichment ought to be provided if this is not being done already.

## IV. USING TEST DATA TO IDENTIFY INSTRUCTIONAL NEEDS

### Interpreting the Individual Item Analysis Report

Exhibit 3 (page 16) illustrates a typical item analysis report. It provides information about each student's performance on each test item. The sample report shows student performance on 10 of the 50 test items from a Math Basic Concepts test. Correct responses are indicated by plus (+) signs, and the percentage of the class (Ms. Thatcher's 6th grade) and the national population answering each question correctly are both reported. Questions are grouped by objective (e.g., place value), or what test publishers frequently refer to as domain.

An example will clarify how information is presented and what it means. Question #5 deals with place value. The answer to this question is "A" -- the correct answer obtained by 11 of the 17 students who attempted the question. Of Ms. Thatcher's class, 65 percent answered this item correctly, whereas 78 percent of the national norms group obtained the correct answer.

### Looking at Group Results

It's usually a good idea to proceed from the general to the specific; from group or classroom data to the performance of individual students.

Using Exhibit 3 to analyze overall group performance, we see that the performance range extends from 41 percent (scores on two properties of operations items) to a high of 76 percent (the score on item #20).

Comparing the columns labeled "This Group" and "National Group/ Percent Choosing Correct Response," we see that for only two of the ten test questions (items 20 and 21) did class performance exceed the performance of the national norms group. Ms. Thatcher's sixth grade is below the national norms on items 5, 7, 16, 46, 40, 6, 12, and 13.

EXHIBIT 3

The Individual Item Analysis Report provides information about each student's performance on each test item. The incorrectly chosen alternative is indicated; the percent in the national population selecting the correct option is also shown. Correct responses are recorded with a +; a blank indicates that the item was omitted. The report is organized by test domains in order to facilitate individual diagnosis.

INDIVIDUAL ITEM ANALYSIS

Report for: MS THATCHER SOUTH ELEMENTARY Grade: 6 National Norms: 6 FALL Number Tested: 17 Form Used: STEPX Test Date: 10/79 Processing Codes: 0 78 20 A 0 78 32 B

| Student Codes | | Student Codes | | Student Codes | | Student Codes | | Student Codes | |
|---|---|---|---|---|---|---|---|---|---|
| 01 | ALLEN BOB L | 06 | BOYD LEE M | 11 | LOGAN PHIL T | 16 | SPANN GEOR L | | |
| 02 | BARR LEE S | 07 | CASE SETH L | 12 | MERTIN DAVI C | 17 | SPEND JILL S | | |
| 03 | BAY MARY N | 08 | DALEY SUE W | 13 | PURCELL DONN G | | | | |
| 04 | BENTON JOE T | 09 | HIRST SAM T | 14 | SMITH LEVIN | | | | |
| 05 | BOGART ANN B | 10 | LEAR MARI J | 15 | SMITH NANG T | | | | |

+ = correct response
A,B,C, or D = wrong response given
Blank = no response given

| | Item Nos. | Correct Response | Student Codes 01 | 02 | 03 | 04 | 05 | 06 | 07 | 08 | 09 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | THIS GROUP Percent Choosing Correct Response | NATIONAL GROUP Percent Choosing Correct Response |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| MATH BASIC CONCEPTS LEVEL H DOMAIN | | | | | | | | | | | | | | | | | | | | | |
| NUMBER AND OPERATION ITEM DESCRIPTIONS | | | | | | | | | | | | | | | | | | | | | |
| PLACE VALUE | 5 | A | + | + | C | + | | B | + | + | + | + | D | + | + | + | | + | B | 65 | 78 |
| | 7 | D | + | A | + | + | + | + | B | | + | + | + | + | | + | C | + | + | 71 | 72 |
| | 16 | A | + | + | + | + | + | | + | B | C | + | | + | + | D | B | C | + | 59 | 60 |
| | 20 | B | + | A | + | + | + | D | + | + | + | + | A | + | + | A | + | + | + | 76 | 70 |
| COMPARING AND ORDERING | 46 | C | B | + | + | + | B | | + | | + | + | + | A | B | + | + | | D | 53 | 58 |
| | 21 | A | + | + | + | B | + | C | + | + | | + | + | + | + | + | + | C | B | 65 | 59 |
| | 40 | D | + | + | B | + | + | A | + | C | + | B | + | + | C | C | + | + | B | 65 | 76 |
| PROPERTIES OF OPERATIONS | 6 | D | + | A | A | | A | + | B | A | + | + | B | B | + | + | | + | | 41 | 48 |
| | 12 | B | A | + | + | C | C | A | + | + | B | + | | + | + | A | C | D | A | 41 | 51 |
| | 13 | A | + | + | D | + | + | | B | + | | + | B | + | C | B | + | + | D | 53 | 60 |



The Uses of Measurement Series. Copyright (c) 1981 Educational Testing Service.

Looking at Domain Results

After briefly scanning overall results, it is generally useful to identify domains on which performance is highest and lowest. Of the three objectives tested (place value, comparing and ordering, and properties of operations), performance is lowest on the three items covering properties of operations.

On the other hand, we notice that performance on the four questions covering place value is considerably better. For the class as a whole, this would suggest that place value does not need reteaching and probably not review at this time, but that properties of operations probably should be retaught. Comparing and ordering might be reviewed, although comparison with performance of the norms group (see last column in Exhibit 3) suggests that two of the three questions in this subset are fairly difficult.

Looking at Individual Results

Before recording individual strengths and weaknesses, it is probably a good idea to ask yourself: "Is this skill or piece of information important for my students to know at this time?" If the skill is important, and students miss questions covering it, you should consider reviewing or reteaching.

There is no hard and fast rule as to when to review or when to reteach, but sometimes performance on a small number of items can raise a flag to alert teachers that some students are having trouble with a particular skill area. For example, returning to Exhibit 3 we see that five students missed at least two of the three items covering comparing and ordering.[1] The skills measured by these items should be reviewed with this group of students.

[1] No response is considered incorrect.

Similarly, we find that three students missed at least three out of the four questions dealing with place value. The teacher should review place value with these students.

Temporary instructional groups might be formed to work on these skill areas, but this organizational solution will not work for all students in the class. From the exhibit we see, for example, that one student, Lee Boyd, has particular difficulty in all three areas. He needs special attention.

Planning Steps

The above discussion suggests a four-step process to use in instructional planning:

1. Examine the overall performance of the class in comparison to the national reference group;

2. Examine clusters of items to determine the range of performance and to identify items on which performance is highest and lowest;

3. Make decisions for the class as a whole;

4. Make decisions for specific students by examining the items to determine which students missed which items.

On the basis of this analysis and other available information, students with common skill needs can be grouped for additional instruction or particular skills can be reviewed with the entire class, as necessary.

## V. WORKING WITH THE PRESS

There are two goals to keep in mind when working with the press.

- Short-term Goals -- You want to get certain information communicated clearly and accurately so that it will be reported correctly.

- Long-term Goals -- You want to build a good relationship with members of the press to improve future efforts at communicating information.

In all of your dealings with the press, be sure that what you do will help you meet both your short-term and your long-term goals. Do not sacrifice a good long-term relationship with the press for the sake of today's story.

### Guidelines for Presenting Information to the Press

- Be Clear -- Test scores can be very complex to understand and interpret fully and accurately. This is true of many other issues as well -- such as school budgets. Do not expect reporters to plow through a lengthy and complex report to dig out the important information. It's the job of the school district to present the information cleanly and directly. You are in a position to go through the information, extract the important facts, and present the information in a clear and straightforward way.

- Be Thorough - Even though reporters may not want to go through a long report to get the information they need, backup information is often necessary, either to follow up on something of interest or to document what they write. Have complete and detailed information available in a usable format that you can supply upon request.

- Be Accurate - If you give out information in such a way that it can be misinterpreted, it probably will be! Make sure your information is correct. This is obviously essential, since you do not want to have to try to get misinformation corrected later. It's also essential for building a good long-term relationship with the press. They need to know that they can trust your facts and figures.

- Be Honest - If test scores fall, do not try to hide it. First of all, you probably cannot. Also, you gain credibility by being honest. If you're willing to openly share the bad news, everyone will be more likely to believe you when you have good news.

- Know Your Data -- You need to be able to answer questions that may arise clearly and accurately. If you hesitate in responding, someone will wonder if you're trying to hide something. Also, if you're not sure of your information, you're more likely to become defensive. And, worst of all, if you're not thoroughly familiar with the data, you may give someone some wrong information.

- Take the Initiative -- Do not wait for the press to come knocking on your door. You need to contact reporters and provide them with whatever information you want reported. Reporters do not want to miss out on a story of interest to the community, so let them know there's information available. Also, they need to know that, in the future, if a story unfolds, you'll let them know about it.

- Be Timely -- Let the press know when certain kinds of information will be available. Also, know and be aware of press time pressures (such as deadlines, printing dates for local papers, and taping times for radio shows). If you want the press to be understanding about the problems you have, be fair to the press by becoming aware of and understanding their problems and constraints.

## Steps to Follow in Working With the Press

There are three stages in communicating information to the press.

### Stage One -- Background

If you find yourself saying, "but they didn't know the situation here," when you read a story about your schools, you need to do something about that. You should take the following three steps to ensure that everyone has the necessary information.

1. Use the checklist which follows to prepare a Background Kit for yourself. You will probably want to update this information periodically, since it may change over time.

- ☐ Names, addresses, and phone numbers of all local and regional newspapers.
- ☐ Names and phone numbers of any reporters who specifically cover education or civic affairs.
- ☐ Deadlines and printing days for newspapers.
- ☐ Names, addresses, and phone numbers of all local or regional radio and television stations.
- ☐ Information on which shows are daily and which are weekly, including taping times and times of live shows.

Having this information compiled ahead of time means that, when you have information to share, you'll be ready to go ahead and contact the appropriate people.

2. Use the checklist which follows to prepare a Background Kit for the press. This information will have to be updated annually so that it remains accurate.

- ☐ Names, addresses, and phone numbers of top school administrators.
- ☐ Names and ways of contacting School Committee members.
- ☐ Information on the district, including:
  --number of pupils
  --number of teachers
  --size of annual budget
  --number of students in special education programs
  --proportions of students in various racial/ethnic groups
  --number of male and female students
  --number of students at each grade level
  --names, addresses, size, and grade levels for each school
- ☐ Information on the testing program, including
  --names of all standardized tests administered by the district
  --approximate annual testing period
  --approximate dates for reporting of results
  --grades tested
  --for Basic Skills tests, the score required to meet district standards
- ☐ Calendar of events, including
  --annual testing dates
  --annual dates when test results are available
  --major dates in the budget preparation process
  --cycle for School Committee elections
  --annual school events (fairs, graduation, etc.)

3. Once your Press Background Kit is prepared and duplicated, it's time to provide the information to the press. You might use the presentation of this Kit as a way of meeting local reporters and broadcasters. For instance, you might call and meet the members of the press individually. Or, you might have a tour of some of your schools and a general information, getting-acquainted session at which time you would distribute the press packet.

   Be sure, though, when you invite the press to this background briefing, to let them know there's no "hot story" happening. Also, take this opportunity to find out what other kinds of information reporters would like to have as background on your district.

You now know your local reporters, and they know something about you and your school district. Now, you are ready for Stage Two.

Stage Two -- Getting Out the News

The test results have just come back, and you're making a formal presentation at next week's School Committee meeting. Its time to talk to the press! Follow these three steps to get the information out to the right people at the right time.

1. Write a press release. Sample press release information is provided in Appendix B. Make sure your press release includes the following information.

   - What are the test results? How does the district compare with other districts in the state? What are the percentile scores for each grade? Where do the percentile ranks fall in relation to the comparison score band?
   - How do the test results compare with last year's results? Is the trend up or down?
   - If the trend is higher or lower, what is the reason? Budget cutbacks? A new reading program? Shorter instructional periods? A redesigned curriculum?
   - What are you going to do with the test results? A new program of some kind? Reallocation of funds? A new task force? If so, when will the public see these things materialize?
   - How do you feel about the test results? If you are pleased, say so. If not, its good to express your concern.

2. Call reporters ahead of time to let them know when and where information will be released. You can provide the press release to reporters ahead of time. The release itself will give the date and time when the news may be released for publication. In other words, if the School Committee is meeting Monday night, you can call reporters in on Friday, give them the press release, and answer their questions. They will not release the results until the time specified on the press release, but having the press release ahead of time allows the reporter to gain a better understanding of the information you are presenting and to ask questions. This early presentation of information is mutually beneficial to your district and to the press.

3. Be sure all reporters have the opportunity to get the press release at the same time. Don't play favorites -- you'll be sorry later,

even if the favored reporter writes a great story this time. If the number of reporters is small, you might call them all and ask if they want to come in and pick up a copy of the press release at a certain time when someone will be available to discuss it with them and answer their questions. Or, you may want to mail out the press release to all reporters at the same time (if you have it written early enough to allow for mail delivery).

Stage Three -- Follow-up

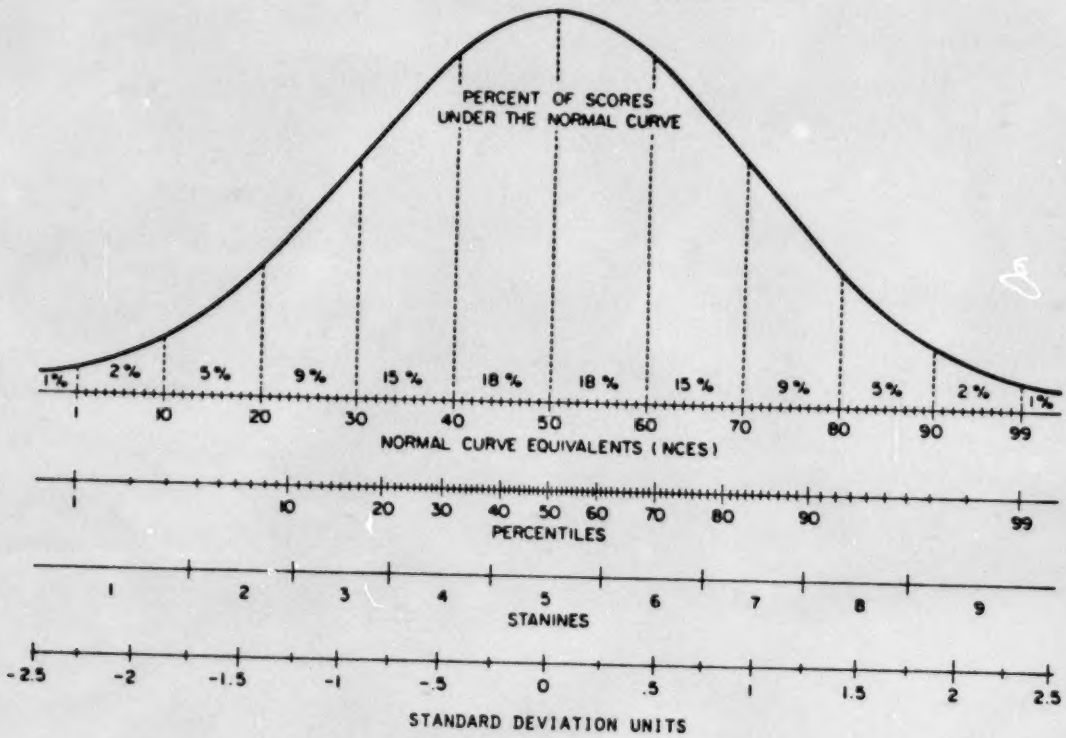
The test results are out, and the press has presented the information clearly and accurately, because you presented it to them clearly and accurately. But your job is not finished!

Make it a point to keep in contact with reporters throughout the year. Let them know what's happening in your district as a result of the test scores. Any changes in the curriculum being planned? What effect will the projected budget have on the program? Any new programs being planned? Any special events happening? Did one of your students or teachers win an award? Let the press know.

These things are not all timely news stories and may or may not get reported. If they do, you're ahead of the game because you have a better informed community. If not, you're still ahead because the members of the press build up knowledge about the system that will help them report the next news story with greater understanding.

## APPENDIX A

AREA UNDER THE NORMAL CURVE DIVIDED INTO NCEs, PERCENTILES, STANINES, AND STANDARD DEVIATION UNITS



## APPENDIX B: SAMPLE PRESS RELEASES

### Sample News Release Outline

Name of district
Title of office sending news release
Name of contact person
Telephone number of contact person
Date of news release

HEADLINE: (Optional)

_____, Superintendent, _____

_____ School District, today announced achievement test results for the district's pupils in grades _____. The tests were administered during the 19__-19__ school year as part of the mandated statewide Basic Skills Improvement Program.

THEN - Summarize the practical significance of test results. Include comparisons over previous years, using quotations from the superintendent.

Basic reading achievement of _____ students in the grades
(district's)
tested (has continued to improve) (has remained constant) during the past few years. _____ said, "_____
(superintendent)
_____.
(quotation on general trend)

NOTE - If one of these two statements cannot be made, start paragraph by breaking down the general trend into grade levels.

Third graders showed the most significant gains in basic reading achievement with _____.

- Indicate any changes in the test(s) used that make certain specific raw scores comparisons difficult.
- Identify the tests used.

- Point out specific areas of instruction and learning measured by this battery of tests, highlighting what is not measured.

--Define "norms" or "standards" and explain how they are determined. Spell out limitations.

--Explain the intent of the testing program.

- If you are pleased with the test results, say so. If not, indicate your concern. Don't become defensive. Above all, be open.

* * * * * * * * *

ADDITIONAL COPY IDEAS:

....Give your staff credit for what they are doing to help raise test scores.

....Identify programs that have been initiated during the current year to accelerate pupil progress.

....Report on special norms (national norms for ESEA Title I pupils, for example) that match grades or schools in your district more closely.

....Stress and document important learning outcomes not measured by standardized tests.

Sample News Release

NEWS RELEASE

Contact: Dr. Susan J. Adams
Superintendent of Schools
Hometown School District
145 Freeman Lane
Hometown, PA 09999

Phone: (999) 222-0000 (office)
(999) 222-3333 (home)

STUDENTS DO WELL IN BASIC SKILLS
For Release May 19, 1984, 12:00 PM

Eighty-eight percent of all students in the Hometown Public Schools met minimum local standards on basic skills tests in reading, writing, and mathematics. The tests, required under the statewide Basic Skills Improvement Policy, were given in April to all students in grades 3, 6, and 8.

For 3rd grade, 90% of students met standards in reading; 86% in mathematics; and 88% in writing.

For 6th grade, 91% of students met standards in reading; 89% in mathematics; and 86% in writing.

For 8th grade, 90% of students met standards in reading; 90% in mathematics; and 84% in writing.

Superintendent of Schools, Dr. Susan J. Adams, said "the results demonstrate that the great majority of Hometown students are competent in all basic skills areas." She attributed the high performance to the development of a district curriculum for basic skills and training opportunities for teachers.

The purpose of the testing is to identify students who are not mastering the basic skills so they can be given help before graduating from high school. According to the district's basic skills plan, developed with the assistance of parents, teachers, administrators, employers and other members of the public, students who did not meet standards are being further diagnosed and will receive additional instruction.

The tests also identify areas in the school program that should be strengthened. According to Robert Jefferson, Associate Superintendent of Schools, the district is planning to review its writing program with special attention to increasing opportunities for students to write. "Even though many of our students are able to write effectively, we believe that greater emphasis in this area will be beneficial," he said.

Students in grades 3 and 6 took the XYZ test. The elementary reading tests measure students' comprehension skills. The elementary mathematics tests measure computation skills. In writing, elementary students were requried to write a paragraph, which was rated on grammar, spelling, usage, and clarity.

Students in grade 8 took the Department of Education's Basic Skills Tests. The state reading test measures students' comprehension skills. The state mathematics test measure students' skills with numbers, computation, estimation, measurement, tables, and statistics. The state writing test requires students to write a letter and an essay, which are rated on mechanics, usage, organization, clarity, and content.

More girls than boys met standards in reading and writing in grades 6 and 8. More boys than girls met standards in mathematics in grade 8. Eighty-eight percent of minority students met reading standards in 3rd grade; 87% at 6th grade; and 91% at 8th grade. In mathematics, 87% of minority students met standards at 3rd grade; 84% at 6th grade; and 92% at 8th grade. In writing, 81% of minority students met standard at 3rd grade; 85% at 6th grade, and 80% met standards at 8th grade.

Superintendent Adams cautioned against using the test results to make comparisons among school systems since "...each community selects its own tests, chooses its own grade levels at which to test, and establishes its own standards on its tests. The major use of the results is to help us provide additional instruction where it is most needed."

The report, which has been submitted to the Department of Education, may be examined at the central office of the school district as well as the main branch of the public library.

BASIC SKILLS IMPROVEMENT POLICY

Department Resources - Assessment

General

| | |
|---|---|
| Information Sources for Selecting or Developing Achievement Tests | February, 1979 |
| Measure for Measure - Evaluating Student Writing | December, 1980 |
| Topic Ideas for Assessment Writing at the Elementary Level | January, 1981 |
| 1980-81 Statewide Summary of Student Achievement of Minimum Standards in the Basic Skills of Reading, Writing and Mathematics | December, 1981 |

Implementation Guides

| | |
|---|---|
| Assessing Basic Skills Achievement - Guide #1 | May, 1980 |
| Writing Assessment Manual - Guide #2 | March, 1981 |
| Standard Setting Manual - Guide #3 | March, 1981 |
| Local Test Development Manual - Guide #5 | November, 1981 |
| Supplement to Writing Assessment Manual | January, 1982 |

Assessment Instruments and Information

| | |
|---|---|
| Massachusetts Test of Basic Skills (Forms 2-6) (Reading, Writing and Mathematics) | 1980 & 1981 |
| Test Development Technical Reports (Forms 2-6) (Reading, Writing and Mathematics) | 1980 & 1981 |
| Massachusetts Assessment of Basic Skills - 1978-79 (Reading, Writing and Mathematics) | September, 1979 |
| Massachusetts Assessment of Basic Skills - 1979-80 (Listening and Speaking) | September, 1980 |
| Assessment of Basic Speaking and Listening Skills: State of the Art and Recommendations for Instrument Development (Summary and Test Reviews) | September, 1979 |
| Development of State Speaking Assessment Instrument: Reliability and Bias Study | June, 1982 |

Most of the above have been distributed to Massachusetts school systems. Readers should check with the appropriate regional center to see if additional copies are available for distribution or examination.